Institute for Advanced Simulation

QM/MM Methodology: Fundamentals, Scope, and Limitations

Walter Thiel

published in

Multiscale Simulation Methods in Molecular Sciences,
J. Grotendorst, N. Attig, S. Blügel, D. Marx (Eds.),
Institute for Advanced Simulation, Forschungszentrum Jülich,
NIC Series, Vol. **42**, ISBN 978-3-9810843-8-2, pp. 203-214, 2009.

http://www.fz-juelich.de/nic-series/volume42

QM/MM Methodology: Fundamentals, Scope, and Limitations

Walter Thiel

Max-Planck-Institut für Kohlenforschung
45470 Mülheim, Germany
E-mail: thiel@kofo.mpg.de

Combined quantum-mechanical/molecular-mechanical (QM/MM) methods have become a popular approach for modeling local electronic events in large systems with thousands of atoms. QM methods are used to describe the active site where chemical reactions or electronic excitations occur, and MM methods are employed to capture the effect of the environment on the active site. This article gives an overview over methodological and practical issues in QM/MM studies and outlines the scope and the limitations of current QM/MM applications.

1 Introduction

The QM/MM concept was introduced in 1976 by Warshel and Levitt who presented the first semiempirical QM/MM model and applied it to an enzymatic reaction[1]. The QM/MM approach found wide acceptance only much later, in the 1990s. Over the past decade, numerous reviews have documented the development of the QM/MM methodology and its application. Here we mention only a few of these[2–8] and refer to our own recent reviews[6,8] for an up-to-date coverage of the field with an extensive literature survey (755 and 627 references, respectively). The reader should consult these reviews for access to the original QM/MM papers since we shall quote only a small selection of these in the following.

The QM/MM approach is by now established as a valuable tool for modeling large biomolecular systems, but it is also often applied to study processes in explicit solvent and to investigate large inorganic/organometallic and solid-state systems. Methodological issues that are common to all these areas will be addressed in Sec. 2, while practical issues and potential pitfalls will be discussed in Sec. 3. Thereafter, an overview over QM/MM applications will be provided in Sec. 4. We conclude with a brief summary in Sec. 5.

2 Methodological Issues

The design of composite theoretical methods gives rise to a number of methodological problems that need to be solved. The basic idea is to retain (as much as possible) the formalism of the methods that are being combined and to introduce well-defined conventions for their coupling. In this section, we address the methodological choices that need to be made in the QM/MM case.

2.1 QM/MM partitioning

The entire system is divided into the inner QM region that is treated quantum-mechanically and the outer MM region that is described by a force field. There is a boundary region at the

interface where the standard QM and MM procedures may be modified or augmented in some way (e.g., by the introduction of link atoms or boundary atoms with special features, see below). The choice of the QM region is usually made by chemical intuition: one can normally define a minimum-size QM region on chemical grounds by considering the chemical problem at hand, and one can then check the sensitivity of the QM/MM results with respect to enlarging the QM region.

Standard QM/MM applications employ a fixed QM/MM partitioning where the boundary between the QM and MM regions is defined once and for all at the outset. It is also possible, but more involved, to allow the boundary to move during the course of a simulation (adaptive partitioning, "hot spot" methods) in order to describe processes with shifting active sites (e.g., the motion of solvated ions)[9].

QM/MM methods can be generalized from two-layer to multi-layer approaches, with a correspondingly extended partitioning. One such example is the use of a continuum solvation model as a third layer to mimic the effects of bulk solvent[10,11]. Other multi-layer approaches such as ONIOM go beyond the original QM/MM concept by integrating two or more QM regions[12].

2.2 Choice of QM method

The selection of a suitable QM method in QM/MM calculations follows the same criteria as in pure QM studies (accuracy and reliability versus computational effort). Traditionally, semiempirical QM methods have been most popular, and they remain important for QM/MM molecular dynamics (MD) where the computational costs are very high. Density functional theory (DFT) is the workhorse in many contemporary QM/MM studies, and correlated ab initio methods are increasingly used in electronically demanding cases or in the quest for high accuracy.

In small-molecule quantum chemistry, one nowadays often attempts to converge the results with regard to QM level and basis set. It has been demonstrated recently that this is also possible in QM/MM work on enzymes: using linear scaling local correlation methods the computed barriers for the rate-determining reactions in chorismate mutase and p-hydroxybenzoate hydroxylase (PHBH) can be converged to within 1–2 kcal/mol at the ab initio coupled cluster LCCSDT(0) level[13,14].

2.3 Choice of MM method

Established MM force fields are available for biomolecular applications (e.g., CHARMM, AMBER, GROMOS, and OPLS) and for explicit solvent studies (e.g., TIP3P or SPC for water). MM methods are generally less developed in other areas such as organometallic or solid-state chemistry which may pose restrictions on corresponding QM/MM work. Even in the favorable biomolecular case, it is often necessary to derive some additional force field parameters (whenever the QM/MM calculations target situations in the active-site region that are not covered by the standard force field parameters).

The classical biomolecular force fields contain bonded terms as well as nonbonded electrostatic and van der Waals interactions. Electrostatics is normally treated using fixed point charges at the MM atoms. The charge distribution in the MM region is thus unpolarizable which may limit the accuracy of the QM/MM results. The logical next step towards

enhanced accuracy should thus be the use of polarizable force fields which are currently developed by several groups in the biomolecular simulation community using various classical models (e.g., induced dipoles, fluctuating charges, or charge-on-spring models). The QM/MM formalism has been adapted to handle polarizable force fields[8,15], but one may expect corresponding large-scale QM/MM applications only after these new force fields are firmly established. In the meantime, essential polarization effects in the active-site environment may be taken into account in QM/MM studies by a suitable extension of the QM region (at increased computational cost, of course).

2.4 Subtractive versus additive QM/MM schemes

Subtractive QM/MM schemes are interpolation procedures. They require (i) an MM calculation of the entire system, (ii) a QM calculation of the inner QM region, and (iii) an MM calculation of the inner QM region. The QM/MM energy is then obtained simply by summing (i) and (ii) and subtracting (iii) to avoid double counting. In such an interpolation scheme, the QM/MM interactions are handled entirely at the MM level. This may be problematic with regard to the electrostatic interactions which will then typically involve fixed atomic charges in the QM and MM regions. Therefore, realistic MM parameters are also needed for the QM region which are often not available and difficult to obtain for typical QM/MM applications (where the QM region is "non-standard" and electronically demanding). These drawbacks have made subtractive QM/MM schemes less attractive, especially in the biomolecular area. On the positive side, it should be noted, however, that subtractive schemes are easy to implement and to generalize to the multi-layer case[12].

Additive schemes require (i) an MM calculation of the outer MM region, (ii) a QM calculation of the inner QM region, and (iii) an explicit treatment of the QM/MM coupling terms. The QM/MM energy is the sum of these three contributions. The coupling terms normally include bonded terms across the QM/MM boundary, nonbonded van der Waals-terms, and electrostatic interaction terms. The former two are generally handled at the MM level (using protocols that avoid double counting and related complications), while the latter one is modeled explicitly. This has the advantage that the electrostatic QM/MM interactions can be described realistically using QM-based treatments (see below). It is probably for this reason that the majority of the currently used QM/MM schemes are of the additive type.

2.5 Electrostatic QM/MM interactions

A hierarchy of models is available for handling the electrostatic coupling between the QM charge density and the MM charge model which may be classified[16] as mechanical embedding (model A), electrostatic embedding (model B), and polarized embedding (models C and D). They differ by the extent of mutual polarization between the QM and MM region.

Mechanical embedding is equivalent to the subtractive QM/MM scheme outlined above in that it treats the electrostatic QM/MM interactions at the MM level (typically between rigid atomic point charges). Both the QM and MM region are unpolarized in this case, and the QM charge density comes from a gas-phase calculation (without MM environment). This will often not be accurate enough, especially in the case of very polar environments (as in most biomolecules).

Electrostatic embedding allows for the polarization of the QM region since the QM calculation is performed in the presence of the MM charge model, typically by including the MM point charges as one-electron terms in the QM Hamiltonian. The electronic structure of the inner region can thus adapt to the environment, and the resulting QM density should be much closer to reality than that from a gas-phase model calculation. The majority of the current QM/MM work employs electrostatic embedding.

Polarized embedding attempts to capture the back-polarization of the MM region by the QM region as well, either in a one-way sense (model C) or in a fully self-consistent manner with mutual polarization (model D). The latter is the most refined embedding scheme which, however, has been applied only rarely up to now. It is expected to become more popular when general-purpose polarizable force fields are being used more often as MM components in QM/MM work, because polarized embedding is the natural coupling scheme in this case. As already mentioned above, polarization effects near the active site can alternatively also be taken into account with standard electrostatic embedding if the QM region is extended accordingly.

2.6 Boundary treatment

In many QM/MM studies it is unavoidable that the QM/MM boundary cuts through a covalent bond. The resulting dangling bond must be capped to satisfy the valency of the QM atom at the frontier, and in the case of electrostatic or polarized embedding, one must prevent overpolarization of the QM density by the MM charges close to the cut. To cope with these problems, there are essentially three different classes of boundary schemes that involve link atoms, special boundary atoms, and localized orbitals, respectively.

Link-atom schemes introduce an additional atomic center (usually a hydrogen atom) that is not part of the real system and is covalently bonded to the QM frontier atom. Each link atom generates three artificial nuclear degrees of freedom that are handled differently by different authors. The most common procedure is to fix the position of the link atom such that it lies in the bond being cut, at some well-defined distance from the QM frontier atom, and to redistribute the forces acting on it to the two atoms of the bond being cut (by applying the chain rule)[17]. This effectively removes the artificial degrees of freedom since the link-atom coordinates are fully determined by the positioning rule rather than being propagated according to the forces acting on them. Concerning the possible overpolarization in link-atom schemes, several protocols have been proposed to mitigate this effect which involve, for example, deleting or redistributing or smearing certain MM charges in the link region. Widely used is the charge-shift protocol[18].

Boundary-atom schemes replace the MM frontier atom by a special boundary atom that participates as an ordinary MM atom in the MM calculation, but also carries QM features to saturate the valency of the QM frontier atom in the QM calculation. These QM features are parametrized such that the boundary atom mimics the cut bond and possibly also the electronic character of the attached MM moiety. Examples for such schemes include the adjusted connection atoms for semiempirical QM methods[19], the pseudobond approach for ab initio and DFT methods[20], and the use of tailored pseudopotentials within plane-wave QM methods[21]. Properly parametrized boundary-atom schemes should be more accurate than link-atom schemes, but they are less popular in practice because the required special parameters are not generally available (only for selected bonds).

Localized-orbital schemes place hybrid orbitals at the boundary and keep some of them frozen such that they do not participate in the SCF iterations. These approaches are theoretically satisfying because they provide a boundary treatment essentially at the QM level. However, they are technically involved (mainly because of the orthogonality constraints that need to be imposed), and require transferability of the localized orbitals between model and real systems. Examples for such schemes are the local SCF method[22] in different variants[8] and the generalized hybrid orbital (GHO) method[23].

There have been several evaluations of and comparisons between the available boundary treatments. Overall the performance of link-atom schemes seems generally on par with localized-orbital approaches: both provide reasonable accuracy when applied with care. In practice, the link-atom scheme is most popular because of its simplicity and robustness, but the GHO treatment is also frequently used.

2.7 QM/MM geometry optimization

In theoretical studies of small molecules, potential energy surfaces (PES) are commonly explored by geometry optimization to locate the relevant stationary points (minima, transition states). This is also possible in QM/MM studies of large molecules with thousands of atoms, in principle, but it is obvious that one needs techniques that can handle thousands of degrees of freedom and are still efficient. The algorithms for manipulating coordinates should ideally be scaling linearly with the number of degrees of freedom, and the optimization should take advantage of the partitioning of the system into a QM region, where energy and gradient evaluation are computationally expensive, and an MM region, where these calculations are almost for free. Among the various approaches that have been proposed in this context[8], we only mention a linear-scaling fragment-based divide-and-conquer optimizer[24] and microiterative optimization strategies[25] with alternating geometry relaxation in the core region (containing the QM region) and the environment. Their combined use allows the efficient optimization of minima and transition states in large molecules at the QM/MM level even when using electrostatic or polarized embedding[26].

Given the vast configuration space that is accessible to the large molecules studied by QM/MM techniques, there are many closely related minima and transition states for any particular chemical reaction. QM/MM geometry optimizations of the stationary points along a single reaction path are therefore of limited significance. It is thus advisable in QM/MM optimization studies to determine at least several representative transition states with their corresponding minima in order to assess the influence of the conformational diversity of the environment; snapshots from classical MD simulations can serve as starting structures. Application of this procedure to the rate-limiting reaction in PHBH has shown that rms fluctuations of the computed QM/MM barriers for 10 snapshots are of the order of 2 kcal/mol[14]. Uncertainties of this magnitude must be anticipated when investigating only a single reaction path.

2.8 QM/MM molecular dynamics

The preceding discussion emphasizes the need for sampling configuration space in large molecules using molecular dynamics or related approaches. QM/MM MD calculations are computationally quite demanding, however, and routinely affordable only at the semiempirical QM/MM level. As in the case of QM/MM geometry optimization, this calls for

special techniques that reduce the computational cost by exploiting the QM/MM partitioning. One strategy is to avoid the expensive direct sampling of the QM region while fully sampling the MM configurations. An early example of this approach[27] kept the QM region fixed while sampling the MM region and used ESP(electrostatic potential)-derived charges for the QM atoms to evaluate the electrostatic QM/MM interactions during the MD run; this was shown to be successful in the context of a QM/MM free energy perturbation treatment in which the entropic contributions from the QM region are estimated separately[27,28]. There are a number of recent other activities to improve the available QM/MM MD technology[7,8].

2.9 QM/MM energy versus free energy calculations

Free energy differences govern chemical thermodynamics and kinetics, and theoretical studies should thus aim at free energy calculations. Statistical mechanics provides various techniques to determine free energy differences through sampling, e.g., thermodynamic integration, umbrella sampling, or free energy perturbation. All these techniques have been used in conjunction with semiempirical QM/MM methods in a straightforward manner[28–30], but they tend to become too expensive with ab initio or DFT QM components. For the latter case, approximate free energy treatments have been devised that have been reviewed recently[7].

In view of the computational effort and the technical difficulties of QM/MM free energy calculations, it is of interest to check how much the QM/MM results for energies and free energies differ in typical cases. There are not yet enough theoretical data available for a systematic assessment. However, judging from the QM/MM energy and free energy barriers for several enzymatic reactions, the differences often appear to be less than 1 kcal/mol for localized chemical events (e.g., hydrogen abstraction in cytochrome P450cam, OH transfer in PHBH, nucleophilic substitution in fluorinase, proton transfer in cystein protease). This confirms that the less demanding QM/MM geometry optimization studies can provide valuable information for many types of reactions.

3 Practical Issues

QM/MM calculations are not yet ”black-box” procedures. Therefore it seems worthwhile to address some of the practical problems and choices that are encountered in QM/MM work.

3.1 QM/MM software

QM/MM applications require efficient programs with wide-ranging functionality. Many of the commonly available QM and MM packages nowadays offer QM/MM capabilities as an add-on. The alternative is a modular approach that links external QM and MM codes via interfaces to a central core which supplies the QM/MM coupling as well as routines for standard tasks such as structure optimization, molecular dynamics, etc. The core also provides a common user interface to the external programs, at least for the most common options. The ChemShell software[31] is an example for such a modular QM/MM implementation which currently supports interfaces to several QM codes (GAUSSIAN,

TURBOMOLE, MOLPRO, ORCA, GAMESS-UK, NWChem, MNDO) and several MM force fields (CHARMM, GROMOS, AMBER, GULP).

When embarking on a QM/MM project it may be easiest to use the QM/MM capability of a standard QM or MM package that one is familiar with. In the long run, modular QM/MM software will offer more flexibility and allow the user to access more combinations of QM and MM methods and, in general, more QM/MM functionality.

3.2 QM/MM setup for biomolecular simulations

QM/MM studies on large systems such as enzymes require realistic starting structures. These will normally be derived from experiment (e.g., X-ray or NMR) because they cannot be generated by purely theoretical means. Small modifications of experimental structures are common in the setup phase, e.g., involving the replacement of an inhibitor by a substrate or the substitution of specific residues to generate the starting structure for a mutant of interest.

The available structural information from experiment is generally not complete and often not error-free. It thus needs to be checked and processed using the protocols that have been developed over the past decades by the classical simulation community. This involves, e.g., adding hydrogen atoms that are missing in X-ray structures, adding water molecules inside the biomolecule in "empty" spots, assigning the protonation states of titrable residues, and checking the orientation of residues in ambiguous cases. The system is then put into a water box and relaxed by a series of constrained energy minimizations and MD runs at the classical force field level; this may necessitate the derivation of force field parameters for the "non-standard" parts of the system. After equilibration, the system is subjected to a classical MD production run from which snapshots are taken as starting geometries for the QM/MM work. These starting structures typically contain the biomolecular system in a droplet of water (normally around 20000–30000 atoms).

It should be emphasized that this setup requires a lot of work prior to the actual QM/MM calculations. Errors and wrong choices (e.g., with regard to protonation states or the number of water molecules near the active site) cannot normally be recovered at a later stage. These issues have been discussed more thoroughly in a previous review[6], and further practical guidance is available in the original papers that deal with these questions[32,33]. Finally, while the preceding considerations have addressed the QM/MM setup for biomolecules, they should apply in an analogous manner to other systems with similar complexity.

3.3 Accuracy of QM/MM results

QM/MM calculations involve a lot of choices (see Sec. 2), and it is therefore very difficult to converge the QM/MM results with regard to all computational options. Typical biomolecular studies may employ DFT/MM calculations with a standard protein force field, electrostatic embedding, and a link-atom boundary treatment with a charge-shift scheme. The latter ingredients are considered as an integral part of the chosen QM/MM approach, and the sensitivity of the QM/MM results with regard to the chosen force field, embedding scheme, and boundary treatment is thus normally not checked (even though the QM/MM results will depend on these choices). On the QM side, different basis sets

are used in most DFT/MM studies to assess basis set convergence, and it is also common practice to check by how much the DFT/MM results change when using a different functional. Given the large computational effort in QM/MM work, it is not too surprising that high-level ab initio QM components are used rather seldom and that systematic convergence studies with respect to QM level and basis set are rare (unlike in small-molecule QM studies).

Conceptually, QM/MM treatments become more realistic upon extension of the QM region because the effects of the QM/MM coupling terms and of the MM force field on the active site should decrease by increasing the distance to the QM/MM boundary. It is thus highly advisable to validate the QM/MM results for any given application through QM/MM test calculations with larger QM regions.

3.4 QM/MM exploration of potential energy surfaces

In QM/MM geometry optimizations of systems with 20000–30000 atoms (see above) it is usually considered sufficient to allow only around 1000 atoms to move (i.e., the active site and the environment within a distance of typically 6–10 Å from the active site) while the outer part of the system remains fixed at the initially prepared snapshot geometry. This convention is beneficial in QM/MM studies of reaction profiles where it is essential to retain the same conformation of the optimized "active" region during the reaction in order to guarantee a smooth reaction path. Experience shows that this requirement can be well satisfied in practice with systems of around 1000 atoms, which becomes progressively more difficult for larger systems. If this requirement is not fulfilled (e.g., by the flip of a distant hydrogen bond or some other remote conformational change), the QM/MM results from geometry optimization become spurious since the PES is no longer smooth[32].

In QM/MM MD simulations of a large biomolecule in a water droplet, the outermost water layer is normally fixed or restrained such that there is no evaporation. Strict convergence criteria need to imposed in the QM part of the calculation to ensure energy conservation during the MD run[29]. Standard procedures can be applied to monitor the convergence of QM/MM MD simulations[29] and to analyze the results[30].

4 Applications

Biomolecular QM/MM studies constitute the largest application area, with enzymatic reactions as the prime target. Our previous reviews list 286 such QM/MM publications between 2001 and early 2006[6], and 179 such papers in the period 2006-2007[8]. A thorough survey of this work is obviously far beyond the scope of this article. Generally speaking, the QM/MM calculations provide detailed mechanistic insight into enzymatic reactions. The QM/MM energy, and particularly the QM/MM interaction energy, can be partitioned into its various components which offers the opportunity to analyze the effect of the protein environment (down to individual residues). Further insights can be gained by comparing the QM/MM results for the complete enzyme with QM results for suitably chosen model systems. In this manner, one can arrive at an improved understanding of the catalytic power of enzymes (as shown, for example, by a recent summary[8] of QM/MM studies on PHBH, chorismate mutase, and cytochrome P450).

QM/MM methods are suitable not only for studying chemical reactions in the active site of a large system, but also for investigating other localized electronic processes such as electronic excitation. In recent years there is an increasing number of QM/MM applications that address spectroscopic properties and electronically excited states. A typical procedure is to perform a DFT/MM geometry optimization or to extract snapshots from a semiempirical QM/MM MD run, followed by single-point calculations of spectroscopic properties at a suitable QM level (with inclusion of the MM point charges of the environment). QM/MM studies of this kind have been performed to compute not only electronic spectra (UV/vis absorption, emission, and fluorescence spectra), but also magnetic resonance spectra (NMR, EPR) and Mössbauer spectra. Examples include color tuning in the UV spectra of rhodopsins[34], NMR chemical shifts in rhodopsins[35] and in vanadium chloroperoxidase[36], as well as EPR and Mössbauer parameters in cytochrome P450cam[37]. QM/MM calculations can also be used to study excited-state reactivity in large systems (e.g., the photoisomerization in photoactive yellow protein[38] or the dynamics of a photoactive C–G base pair in DNA[39]).

Another QM/MM application area is experimental structure refinement of large biomolecular systems. The basic idea is to use a QM/MM, rather than a pure MM, model that is refined against the experimental data[40]. This is particularly advantageous in and around the active site since the standard biomolecular force fields are less reliable for the inhibitors or substrates that are present in this region. This approach has been applied to the refinement of X-ray, NMR, and EXAFS data[8].

The QM/MM applications outlined so far have been concerned with large biomolecules. As mentioned in the Introduction, QM/MM methods have also often been used to study processes in explicit solvent and in inorganic/organometallic and solid-state chemistry. An overview over these activities is beyond the scope of this article, leading references are available in our recent review[8].

5 Concluding Remarks

QM/MM methods are by now established as a powerful computational technique to treat reactive and other electronic processes in large systems. They can be applied whenever one needs to model a localized electronic event in an active site (typically of the order of 100 atoms) that is influenced by an interacting larger environment. Since they are not yet ”black-box” methods, one should exercise great care in the choice of the various computational QM/MM options and in the assessment of the results obtained. Despite the need to improve the available QM/MM tools further, especially with regard to higher accuracy and better sampling, there is a growing number of successful QM/MM applications in all branches of chemistry. This indicates that the existing QM/MM methods are good enough for the realistic modeling of real-world chemical problems.

Acknowledgments

This research was supported by the Max Planck Society. Many coworkers made essential contributions to our own QM/MM studies that have been mentioned in this article. Their names are listed in the references.

References

1. A. Warshel and M. Levitt, *Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme*, J. Mol. Biol. **103**, 227–249, 1976.
2. J. Gao, *Methods and applications of combined quantum mechanical and molecular mechanical potentials* in Reviews in Computational Chemistry, K. B. Lipkowitz and D. B. Boyd (Eds.), VCH, New York, Vol. **7**, 119-185, 1996.
3. P. A. Sherwood, *Hybrid quantum mechanics/molecular mechanics approaches* in Modern Methods and Algorithms of Quantum Chemistry, J. Grotendorst (Ed.), John von Neumann Institute for Computing, NIC Series Vol. **3**, 285-305, 2000.
4. J. L. Gao and D. G. Truhlar, *Quantum mechanical methods for enzyme kinetics*, Annu. Rev. Phys. Chem. **53**, 467–505, 2002.
5. R. A. Friesner and V. Guallar, *Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics methods for studying enzymatic catalysis*, Annu. Rev. Phys. Chem. **56**, 389–427, 2005.
6. H. M. Senn and W. Thiel, *QM/MM methods for biological systems* in Topics in Current Chemistry, M. Reiher (Ed.), Springer, Berlin, Vol. **268**, 173-290, 2007.
7. H. Hu and W. Yang, *Free energies of chemical reactions in solution and in enzymes with ab initio quantum mechanics/molecular mechanics methods*, Annu. Rev. Phys. Chem. **59**, 573–601, 2008.
8. H. M. Senn and W. Thiel, *QM/MM methods for biomolecular systems*, Angew. Chem. Int. Ed. **48**, 1198–1229, 2009.
9. A. Heyden, H. Lin, and D. G. Truhlar, *Adaptive partitioning in combined quantum mechanical and molecular mechanical calculations of potential energy functions for multiscale simulations*, J. Phys. Chem. B **111**, 2231–2241, 2007.
10. P. Schaefer, D. Riccardi, and Q. Cui, *Reliable treatment of electrostatics in combined QM/MM simulation of macromolecules*, J. Chem. Phys. **123**, 014905/1–14, 2005.
11. T. Benighaus and W. Thiel, *Efficiency and accuracy of the generalized solvent boundary potential for hybrid QM/MM simulations: Implementation for semiempirical Hamiltonians*, J. Chem. Theory Comput. **4**, 1600–1609, 2008.
12. M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber, and K. Morokuma, *ONIOM: A multilayered integrated MO + MM method for geometry optimizations and single point energy predictions*, J. Phys. Chem. **100**, 19357–19363, 1996.
13. F. Claeyssens, J. N. Harvey, F. R. Manby, R. A. Mata, A. J. Mulholland, K. E. Ranaghan, M. Schütz, S. Thiel, W. Thiel, and H.-J. Werner, *High accuracy computation of reaction barriers in enzymes*, Angew. Chem. Int. Ed. **45**, 6856–6859, 2006.
14. R. A. Mata, H.-J. Werner, S. Thiel, and W. Thiel, *Toward accurate barriers for enzymatic reactions: QM/MM case study on p-hydroxybenzoate hydroxylase*, J. Chem. Phys. **128**, 025104/1–8, 2008.
15. D. P. Geerke, S. Thiel, W. Thiel, and W. F. van Gunsteren, *Combined QM/MM molecular dynamics study on a condensed-phase S_N2 reaction at nitrogen: The effect of explicitly including solvent polarization*, J. Chem. Theory Comp. **3**, 1499–1509, 2007.
16. D. Bakowies and W. Thiel, *Hybrid models for combined quantum mechanical and*

molecular mechanical approaches, J. Phys. Chem. **100**, 10580–10594, 1996.
17. U. Eichler, C. M. Kölmel, and J. Sauer, *Combining ab initio techniques with analytical potential functions for structure predictions of large systems: Method and application to crystalline silica polymorphs*, J. Comp. Chem. **18**, 463–477, 1997.
18. P. Sherwood, A. H. de Vries, S. J. Collins, S. P. Greatbanks, N. A. Burton, M. A. Vincent, and I. H. Hillier, *Computer simulation of zeolite structure and reactivity using embedded cluster methods*, Faraday Discuss. **106**, 79–92, 1997.
19. I. Antes and W. Thiel, *Adjusted connection atoms for combined quantum mechanical and molecular mechanical methods*, J. Phys. Chem. A **103**, 9290–9295, 1999.
20. Y. Zhang, T.-S. Lee, and W. Yang, *A pseudobond approach to combining quantum mechanical and molecular mechanical methods*, J. Chem. Phys. **110**, 46–54, 1999.
21. A. Laio, J. van de Vondele, and U. Rothlisberger, *A Hamiltonian electrostatic coupling scheme for hybrid Car-Parrinello molecular dynamics simulations*, J. Chem. Phys. **116**, 6941–6947, 2002.
22. V. Thery, D. Rinaldi, J. L. Rivail, B. Maigret, and G. G. Ferenczy, *Quantum mechanical computations on very large molecular systems: The local self-consistent field method*, J. Comp. Chem. **15**, 269–282, 1994.
23. J. Gao, P. Amara, C. Alhambra, and M. J. Field, *A generalized hybrid orbital (GHO) method for the treatment of boundary atoms in combined QM/MM calculations*, J. Phys. Chem. A **102**, 4714–4721, 1998.
24. S. R. Billeter, A. J. Turner, and W. Thiel, *Linear scaling geometry optimisation and transition state search in hybrid delocalized internal coordinates*, Phys. Chem. Chem. Phys. **2**, 2177–2186, 2000.
25. F. Maseras and K. Morokuma, *IMOMM: A new integrated ab initio + molecular mechanics geometry optimization scheme of equilibrium structures and transition states*, J. Comp. Chem. **16**, 1170–1179, 1995.
26. J. Kästner, S. Thiel, H. M. Senn, P. Sherwood, and W. Thiel, *Exploiting QM/MM capabilities in geometry optimization: A microiterative approach using electrostatic embedding*, J. Chem. Theory Comput. **3**, 1064–1072, 2007.
27. Y. Zhang, H. Liu, and W. Yang, *Free energy calculation on enzyme reactions with an efficient iterative procedure to determine minimum energy paths on a combined ab initio QM/MM potential energy surface*, J. Chem. Phys. **112**, 3483–3492, 2000.
28. J. Kästner, H. M. Senn, S. Thiel, N. Otte, and W. Thiel, *QM/MM free-energy perturbation compared to thermodynamic integration and umbrella sampling: Application to an enzymatic reaction*, J. Chem. Theory Comput. **2**, 452–461, 2006.
29. H. M. Senn, S. Thiel, and W. Thiel, *Enzymatic hydroxylation in p-hydroxybenzonate hydroxylase: A case study for QM/MM molecular dynamics*, J. Chem. Theory Comput. **1**, 494–505, 2005.
30. J. Kästner and W. Thiel, *Bridging the gap between thermodynamic integration and umbrella sampling provides a novel analysis method: Umbrella integration*, J. Chem. Phys. **123**, 144105/1–5, 2005.
31. `http://www.chemshell.org`
32. A. Altun, S. Shaik, and W. Thiel, *Systematic QM/MM investigation of factors that affect the cytochrome P450cam-catalyzed hydrogen abstraction of camphor*, J. Comp. Chem. **27**, 1324–1337, 2006.
33. J. Zheng, A. Altun, and W. Thiel, *Common system setup for the entire catalytic cy-*

cle of cytochrome P450cam in quantum mechanical/molecular mechnical studies, J. Comp. Chem. **28**, 2147–2158, 2007.
34. M. Hoffmann, M. Wanko, P. Strodel, P. H. König, T. Frauenheim, K. Schulten, W. Thiel, E. Tajkhorshid, and M. Elstner, *Color tuning in rhodopsins: The mechanism for the spectral shift between bacteriorhodopsin and sensory rhodopsin II*, J. Am. Chem. Soc. **128**, 10818–10828, 2006.
35. J. A. Gascon, E. M. Sproviero, and V. S. Batista, *Computational studies of the primary phototransduction event in visual rhodopsin*, Acc. Chem. Res. **39**, 184–193, 2006.
36. M. P. Waller, M. Bühl, K. R. Geethalakshmi, D. Wang, and W. Thiel, *Vanadium NMR chemical shifts calculated from QM/MM models of vanadium chloroperoxidase*, Chem. Eur. J. **13**, 4723–4732, 2007.
37. J. C. Schöneboom, F. Neese, and W. Thiel, *Towards identification of the Compound I reactive intermediate in cytochrome P450 chemistry: A QM/MM study of its EPR and Mössbauer parameters*, J. Am. Chem. Soc. **127**, 5840–5853, 2005.
38. G. Groenhof, M. Bouxin-Cademartory, B. Hess, S. P. de Visser, H. J. C. Berendsen, M. Olivucci, A. E. Mark, and M. A. Robb, *Photoactivation of the photoactive yellow protein: Why photon absorption triggers a trans-to-cis isomerization of the chromophore in the protein*, J. Am. Chem. Soc. **126**, 4228–4233, 2004.
39. G. Groenhof, L. V. Schäefer, M. Boggio-Pasqua, M. Goette, H. Grubmüller, and M. A. Robb, *Ultrafast deactivation of an excited cytosine-guanine base pair in DNA*, J. Am. Chem. Soc. **129**, 6812–6819, 2007.
40. U. Ryde, L. Olsen, and K. Nilsson, *Quantum chemical geometry optimizations in proteins using crystallographic raw data*, J. Comp. Chem. **23**, 1058–1070, 2002.